26 | 系统集成：为什么最容易出问题的是系统集成？

2020-07-15 臧萌

职场求生攻略 进入课程 >

讲述：臧萌
时长 11:13 大小 10.29M

你好，我是臧萌。在上一节中我们聊了软件系统架构师。软件架构师完成了最核心的系统，但是这还远远没完。如果你想让这个系统真的运转起来，那还需要系统集成。而且在我看来，系统集成很多时候所占的时间和精力，远比系统核心要多。

那到底什么是系统集成呢？系统集成，简单来说，就是要对接、落地，验证，看到真正的成果。

吃上一道新菜，过程远比你想的复杂

我们拿饭店来打个比方。假设我们现在运营着一个连锁饭店，中心厨房会为每个连锁饭店配送半成品的菜品。现在我们要推出一款叫“三菌笋片烩鸡丁”的新菜品。对于新菜品来

说，一开始我们要先构思，然后去小范围地验证这个菜品。如果在这个小范围内得到一致好评，那第一步就算完事了。

这时候，如果用软件系统研发的进度做对比的话，菜品研发部分，就是完成了系统架构和关键模块的编写，也就是完成了核心系统的开发。请大家试吃就是搞出一个演示给大家看，结果就是大家看着核心系统的演示，都觉得功能很棒，翘首期待能早日用上。

这是决定这个系统特色的一步，也是最有创造性的一步。但得到了大家的认可，这能算成功了吗？远远没有。这只能算是迈出了通向成功的第一步，距离成功，还差着系统集成这关键的一步呢。

回到我们的例子。我们的目的是为连锁饭店推出一道新的菜品。我们需要给连锁饭店的食客们安全快速地呈上这道新菜，才能算成功。而一个菜能在厨房做出来，和一个菜能一年能在所有饭店长期保持稳定水准，是完全两个概念，难度也是两个级别。

三菌笋片烩鸡丁这道菜，用到了三种菌菇以及笋片和鸡丁。三种菌菇可能无法保证每一个连锁店所在地都有货，可能某些菌菇即使有货，价格也波动很大，可能有的菌菇对保鲜和运输的要求比较高，弄不好就容易变质，造成食品安全问题。

饭店一旦要连锁，类似的问题就会不断地暴露出来，并且不断放大。如果类似问题处理不好，那么这个菜品就是失败的。而这个让新菜品落地的过程，类比到软件系统里，就是系统集成。

如何不掉进软件系统集成那些坑

那我们继续回到软件系统。软件系统在特定的场景，无其他干扰因素的理想情况下正常工作，并不代表它就可以用了。外界的很多因素，都有可能影响系统发挥出本身的功能。系统集成的过程，才是对系统的大考。

每个系统面对的情况都不一样，这里我还是用一个电商网站来举例子，给大家聊聊软件系统集成中常见的那些坑、如何尽量避免掉进坑里、以及要是你不小心掉进去的话，如何快速而优雅地爬出来。

电商网站有库存、商品展示、订单管理、退换货等等子系统。退换货子系统需要对接供货商的 ERP 系统。

系统内部集成：架构统一

所谓系统内部的集成，就是一个系统各个模块之间的集成。比如将库存记录、进出库存、库存盘点等模块等，集成为一个完整的库存系统。如果一个系统设计得过关，那么对于这种集成来说，出问题的几率还是比较小的，而且即使有问题，也容易解决。原因在于这个系统本身就是新的，是遵循着一个统一的架构设计做出来的。各个模块之间在设计之初，就已经做好了集成的准备。

这时候出现的问题，大多是因为一些实现的细节，或者是设计时没有描述清楚造成的。比如代码接口定义不统一，数据交互的格式不统一等等。我之所以说这时候还比较容易解决，就是因为新系统没有历史负担，不需要考虑兼容性等问题，所以只要大家将之前的误解消除，解决问题的工作量还是可以估算的。

这时候如果出现比较大的问题，很可能是因为模块由不同的架构师来负责，而不同的架构师之间没有充分的沟通。好的方式，应该是让一个架构师负责一个完整的模块，根据开闭原则，提供统一且简单清晰的接口，对外提供功能。

正所谓，令出多门，兵家大忌。如果很多人都可以拍板做决定，那么事情到头来很可能会乱套。即使需要不同的架构师负责，也要有一个总的架构师，提供统一的总体架构，解决架构师之间的争论。

好的，这时候我们有了一个库存模块，功能都测试过了，而且它接口简单精炼，一张 A4 纸就能打印出来。别的模块也纷纷告捷，大家满心欢喜地开始向前推进。

外部系统集成：不做假设

这时候，退换货模块需要就退货功能，跟外部 ERP 系统集成了。当你自信满满地问：你们的退货接口是什么，发给我看看。对方发给你一个文档，描述了退货需要提供的数据。你看完，发现需要的数据没啥特殊的，一切尽在掌握。

于是你按照要求，对接的代码就写了起来，没两天就弄好了，和对方说：你们的 RESTful 接口给我测试一下，我这边的数据弄好了。对方一脸问号地说，我们没有什么 RESTful 接口，你把数据保存成那个格式，上传到我们这个 FTP 就行，我们三天内处理完，然后你再去那个 FTP 上，下载一个叫 yyyy-mm-dd-result.json 的文件就可以了，很简单的。

这时候，你终于知道了什么叫做两个世界的人，在你眼里，对方重新定义了简单。同时，你也体会到了什么叫做失控。按照之前的设计，退换货是需要即时给出结果的。这种通过 FTP 传递数据，需要三天处理的模式，完全打破了之前的设计。且不说很多内部的功能都要重新设计重新做，对外的接口也要大改，而且还会影响所有已经和退换货模块集成好的模块。

和外部系统集成，越早确定细节越好，不要假设任何东西。任何在你看来很自然的东西，可能对方就是没有。你的系统如果对这些想当然的东西有强依赖，那么代价可能是惨重的。

这里我可以再举一个例子。如果你的系统的核心功能依赖于某个外部系统的功能，比如某个数据库功能，那你就先把这个通路确定好。比如公司购买的数据库是否有这个功能、公司的 DBA 是否允许我们使用这个功能、这个功能的吞吐量有多大等等。

如果在集成的时候发现，你的系统核心功能与外部系统依赖之间存在问题，这会变得非常麻烦。因为外部系统已经存在，外部系统能不能因为你的需要来做出改变，是非常不确定的。而不同的系统之间，本来就没有统一的架构，系统之间的集成点，需要在设计之初反反复复确认并验证。

漫漫集成路之问

下面我们来谈谈系统集成中其它常见问题。

首先就是配置。有些模块的配置非常繁琐，甚至配置比代码都麻烦。有些系统的配置需要经验的积累和专业的学习，比如 Kafka，HBase，Hystrix 等中间件系统。和这些模块或系统集成，代码之外的配置将是一个非常容易出问题的点，比如说，配置不合理、配置被覆盖、配置非法导致没有生效等等。

如何应对这类集成呢？我的建议是眼见为实。在你用到配置的地方，将配置用 log 的形式输出出来，方便用二分法确定问题。

接着是数据。其实系统集成的时候，很大一部分问题是和数据相关的。这里我给出两类常见的数据问题，一个是用户输入的数据，一个是数据交换。

对于用户输入的数据，简单来说，一个字节都不要相信。用户可以给用户名填写 100k 个字符、用户填写的日期可能是未来或者几百年前、用户可能输入任何数据。所以在使用用户的数据之前，要把数据全部检查一遍。还是那句话，不要做任何假设，哪怕前端已经做了数据验证，自己的系统也要保护性地验证一下数据。

至于数据交换，也是一个非常容易出问题的点。系统集成少不了数据交换，因为大家对数据 schema 定义的不同，数据交换往往伴随着数据格式的转换等操作。这时候，不同系统对于数据缺失、默认值、数据约束等规定的不同，可能会造成各种问题，甚至出错。数据在不同系统之间，交换还可能伴随着数据丢失等问题。

所以，外部系统给的数据，一个字节也不要相信。当和外部系统进行数据交换的时候，一定要在系统交互的边界，好好记 log。要能够将数据的输入、转换、出错等各种情况都记录详细的 log，方便排查问题。

总结

系统集成为什么这么麻烦，这么漫长，原因在于一方面，它需要给出最终的被大家认可的成果。另一方面，在现实条件中可能有各种各样，之前没有想到的问题。这些问题，都会在系统集成的时候，一一暴露出来，成为系统集成路上的拦路虎，造成无法交付成果的情况。

每当我听到有同学说：我代码写完了，这个事情可以结束了。我都会露出尴尬而不失礼貌的微笑。我不能否定这个结论，因为如果你问我还有什么事情没弄完，我也可能说不出个 123 来。但是我敢肯定，这位同学可能还要付出一倍甚至数倍写代码的时间，才能让这个代码真的跑起来，才能说这个事情结束了。

这个过程，就是系统集成，是我们在为自己的想当然、对需求的理解偏差、经验的欠缺、能力的不足、视野的限制、不充足的前期调研、未预料到的意外来付出的代价。功力深厚的架构师，也是一个系统集成的大师。他可以在设计阶段，根据自己的经验，尽量减少问题，降低难度。所以，能够越少地付出这个代价，越能代表自身综合能力的提升，也是代表自身价值的提升。

思考题

系统集成，是任何一个系统都逃不过的大考。你有过哪些关于系统集成刻骨铭心的经历呢？关于系统集成，你遇到过哪些问题，又有哪些实践经验可以分享呢？

好，今天的课程就结束了，希望可以帮助到你，也希望你在下方的留言区和我参与讨论，并把文章分享给你的朋友或者同事，一起交流一下。

提建议

更多课程推荐

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 25 | 系统架构：如何从写代码的程序员，成长为软件系统架构师？

下一篇 27 | 答疑篇：什么样的技术观能够更快成长？

精选留言 (5)

写留言

Sdylan
2020-07-15

现在模式：内部系统-接入模块-外围系统。涉及到防火墙是否开通、如何对接接入模块以及外围系统要做啥。对接外部系统，真的要把上送的参数、返回的结果打印出来。毕竟接入模块不一定是自己团队的，他负责转发而已。

展开

作者回复: 没log，出了问题就是一脸懵逼。

而系统集成一定会出问题的（手动狗头）

2

Middleware
2020-07-15

和外部第三方对接，千万记录好log ，否则都是泪

展开

作者回复: 0假设，外部数据一个字节都不要信

2

Geek_3b1096
2020-07-15

低估集成所需的时间

展开

作者回复: 如果不知道需要多久就把开发的时间 x 2，如果乐观一点就 x 1。

1

再来二两杜康酒
2020-07-16

集成是偷不得懒的，项目排计划也是。当我听到组里人说做完了，心想是去验证的时候了。

作者回复: 哈哈，是这么个意思🤝

南湾小猪
2020-07-16

一位前辈曾分享过做system integration的经验是”Never trust a client!”换句话说，就是对传入进来的数据要多验证。

展开

作者回复: deideidei